THE PROBABILITY THAT A RANDOM MONIC p-ADIC POLYNOMIAL SPLITS INTO LINEAR FACTORS

JOE BUHLER, DANIEL GOLDSTEIN, DAVID MOEWS, AND JOEL ROSENBERG

Contents

Abstract. Let n be a positive integer and let p be a prime. We calculate the probability that a random monic polynomial of degree n with coefficients in the ring $\mathbf{Z}_p$ of p-adic integers splits over $\mathbf{Z}_p$ into linear factors.

1. Introduction

Let R be a complete discrete valuation ring with finite residue field. The main result of this note, Theorem 1.1, is a formula, for each positive integer n, for the probability that a random monic polynomial of degree n with coefficients in R splits over R into linear factors.

By a standard result, R is either the ring of formal power series over a finite field, or the ring of integers of a finite extension of the field of p-adic numbers for some prime p. In either case we write k for the residue field of R and set $q = |k|$. We encourage the reader to focus on the special cases $R = \mathbf{Z}_p$ and $GF(p)[[t]]$, as the statements and proofs in the general case are essentially the same as in these two special cases.

Given $(c_1, \ldots, c_n)$ in R^n, we form the monic (i.e. leading coefficient 1) polynomial $f(x) = x^n + c_1x^{n-1} + \cdots + c_n$. Let $P_n(R)$ be the set of monic degree n polynomials with coefficients in R, and identify $P_n(R)$ with R^n as above. Let $S_n(R) \subseteq P_n(R)$

Date: February 8, 2008.
1991 *Mathematics Subject Classification.* 11S05.
Key words and phrases. p-adic, polynomial, splits completely.

consist of those polynomials that split over R into linear factors (or equivalently, have n roots in R, counting multiplicity). We identify $S_n(R)$ with a subset of R^n. In Section 3 we will similarly write $P_n(k)$ (resp. $S_n(k)$) for the set of monic polynomials of degree n (resp. monic polynomials of degree n that split completely) over the residue field k.

The compact abelian group R has a Haar measure μ of total mass 1. Give R^n the product topology and product measure μ_n (alternatively, view R^n as a compact abelian group in its own right). One can show that $S_n(R)$ is μ_n-measurable as follows. Given $(a_1, \dots, a_n) \in R^n$, we can form the polynomial $\prod_{1 \le i \le n}(x - a_i)$. The resulting map Ξ from R^n to $P_n(R)$ is continuous and has image $S_n(R)$. Thus, $S_n(R)$ is compact, being the continuous image of a compact set, whence closed and in particular measurable.

A main result of this note is a recursive formula for $r_n := \mu_n(S_n(R))$. Loosely speaking, this is the probability that a random monic polynomial of degree n over R splits completely over R. Let q be the order of the residue field of R.

The first four values of r_n are:

$$\begin{aligned} r_1 &= 1, \\ r_2 &= q/2(q+1), \\ r_3 &= (q^2 - q + 1)(q-1)q^3/6(q+1)(q^5-1), \\ r_4 &= h(q)(q-1)^4 q^6/24(q^2-1)^2(q^5-1)(q^9-1), \end{aligned}$$

where $h(q) = q^8 - 2q^7 + q^6 + 2q^5 - q^4 + 2q^3 + q^2 - 2q + 1$.

We also give a formula for the probability that a random polynomial that is not necessarily monic of degree n splits over R.

A student of the first-named author, Asher Auel [1], working independently, has calculated the Jacobian determinant of the map Ξ, then calculated the resulting n-fold p-adic integral. His approach offers an alternative to the method of this note.

Here is our main result. Set $r_0 = 1$. For a sequence $d = (d_0, d_1, \dots, d_{q-1})$ of nonnegative integers, write $|d|$ for their sum, $d_0 + d_1 + \dots + d_{q-1}$.

Theorem 1.1. *Let R be a complete discrete valuation ring with finite residue field k. Set $q = |k|$. For n a positive integer, write r_n for the probability that a random degree n monic polynomial over R splits into linear factors, and set $r_0 = 1$. If $n \ge 0$, then*

$$r_n = \sum_{|d|=n} \prod_{0 \le i \le q-1} q^{-\binom{d_i+1}{2}} r_{d_i}. \tag{1}$$

Remark 1.2. Note that r_n appears in q terms on the right, each time with coefficient $q^{-\binom{n+1}{2}}$. But $0 = 1 - \binom{n+1}{2}$ if and only if $n = 1$. So, as long as $n > 1$, we can solve for r_n in terms of $r_0, \dots, r_{n-1}$.

Here is an approach to r_n using power series. We define $F(t)$ and $G(t)$ in $\mathbf{Q}[[t]]$ by the formulas

$$F(t) = \sum_{n \ge 0} r_n t^n, \text{ and} \tag{2}$$

$$G(t) = \sum_{n \ge 0} \frac{r_n}{q^{\binom{n+1}{2}}} t^n. \tag{3}$$

Then Theorem 1.1 gives that

$$F = G^q. \tag{4}$$

We use a method due to Euler to get an efficient method for computing r_n.

Corollary 1.3. *For $n \geq 0$, we have*

$$\sum_{0 \leq j \leq n} (n - (q+1)j) r_{n-j} \frac{r_j}{q^{\binom{j+1}{2}}} = 0.$$

In particular, r_n is a rational function in q.

2. Proofs

For $r \in R$, consider the map ψ_r that takes the monic degree n polynomial $f(x)$ to $f(x-r)$. Thus, ψ_r maps $P_n(R)$ to itself.

Lemma 2.1. *Let $r \in R$. The map ψ_r from $P_n(R)$ to itself preserves the measure μ_n.*

Proof. Our identification of $P_n(R)$ with R^n makes ψ_r into an affine map. For, write $f = x^n + f_1$ with $\deg(f_1) < n$. Then ψ_r is the composition of the linear map $\psi'_r(x^n + f_1) = x^n + f_1(x-r)$ and the translation map $f \mapsto f + (x-r)^n - x^n$.

Since any translation preserves measure, it is enough to check the linear map ψ'_r.

A basis for $P_n(R)$ is given by the set of monomials x^i for $0 \leq i \leq n-1$. In terms of this basis, the matrix of ψ'_r is upper triangular with ones on the diagonal. Hence it has Jacobian determinant equal to 1. □

We proceed by a series of lemmas. The basic idea is as follows. Suppose the monic polynomial $f(x)$ in $P_n(R)$ splits completely over R. Write πR for the maximal ideal of R, and write $\overline{f}$ for the reduction of f mod πR. Then $\overline{f}$ splits completely over $k = R/\pi R$. Write this factorization as

$$\overline{f}(x) = \rho(x) = \prod_{\alpha \in k} \rho_\alpha(x), \qquad \rho_\alpha(x) = (x-\alpha)^{d_\alpha}. \tag{5}$$

The polynomials $\rho_\alpha(x) = (x-\alpha)^{d_\alpha}$ are pairwise relatively prime. Hence by Hensel's lemma and induction, $f = \prod_{\alpha \in k} f_\alpha$, where f_α in $R[x]$ reduces to ρ_α in $k[x]$.

Of course, the probability that f reduces to ρ (or to any given degree n monic polynomial in $k[x]$) is exactly q^{-n}. The desired result will follow from the following lemma.

Lemma 2.2.

1. $\mathrm{pr}(f \text{ splits completely} | \overline{f} = \rho) = \prod_{\alpha \in k} \mathrm{pr}(f_\alpha \text{ splits completely} | \overline{f_\alpha} = \rho_\alpha)$.
2. $\mathrm{pr}(f_\alpha \text{ splits completely} | \overline{f_\alpha} = \rho_\alpha) = q^{-\binom{d_\alpha}{2}} r_{d_\alpha}$.

To prove Theorem 1.1 from the lemma, it suffices to choose a bijection from k to the set $S = \{0, 1, \ldots, q-1\}$ and to use that $d_i = \binom{d_i+1}{2} - \binom{d_i}{2}$ and that $|d| = \sum_{i \in S} d_i$.

Let g be a polynomial in $P_n(k)$. Write P_g for the set of polynomials f in $P_n(R)$ whose reduction $\overline{f}$ is g. As P_g is a measurable subset of $P_n(R)$, it inherits a topology from $P_n(R)$, and also a measure μ_n, which we renormalize so that P_g has total measure 1. Part (1) of Lemma 2.2 now follows from:

Lemma 2.3. *(Measure theoretic version of Hensel's lemma). Let g and h in $k[x]$ be relatively prime monic polynomials of degree m and n, respectively. The multiplication map*

$$M : P_g \times P_h \to P_{gh}$$

is a bijection, in fact an isomorphism of topological spaces, by the usual Hensel's lemma. Then M preserves measure, that is, $M_*(\mu_m \times \mu_n) = \mu_{m+n}$.

The lemma is an immediate consequence of the following, somewhat more general result.

Lemma 2.4. *Let $A = \prod_{1\leq i} a_i$ and $B = \prod_{1\leq i} b_i$ be countable products of finite sets. Normalize counting measure so that each of a_i and b_i has total mass 1, for all i, and give A and B the product measure. Suppose there is a compatible system of bijections ϕ_n between the partial products $A_n = \prod_{1\leq i\leq n} a_i$ and $B_n = \prod_{1\leq i\leq n} b_i$. That is, for each $n \geq 1$, the map ϕ_n from A_n to B_n is bijective, and for $m \leq n$, if y in A_n has x in A_m as initial string, then $\phi_n(y)$ has $\phi_m(x)$ as initial string. Then the map $\phi = \varprojlim \phi_n$ from A to B is a measure-preserving bijection.*

Proof. Certainly ϕ is bijective. Since ϕ takes basic open sets to basic open sets of the same volume, it follows that ϕ preserves measure. □

Proof of part (2) of Lemma 2.2. By Lemma 2.1, it suffices to treat the case $\alpha = 0$. Set $f = f_0$. We have f in $R[x]$ monic with reduction x^n in $k[x]$. Assume for the moment that $f = x^n + c_1x^{n-1} + \cdots + c_n$ does split completely, say as

$$f = (x - a_1)\cdots(x - a_n).$$

A root of f reduces to a root of $\overline{f}$, so that a_i lies in πR for each i, whence a necessary condition for f to split completely is: c_i lies in $\pi^i R$ for each i. The probability that this condition holds is $q^{-\binom{n}{2}}$, the product of q^{-i} for $0 \leq i \leq n-1$. Set

$$\tilde{f} = f(\pi x)/\pi^n.$$

The necessary condition is met if and only if $\tilde{f}$ has coefficients in R. Conditioned on its being met, $\tilde{f}$ is distributed like a random polynomial in $P_n(R)$, and the result follows, because f splits completely if and only if $\tilde{f}$ splits completely. □

Corollary 1.3 follows from the following efficient recursive computation of the coefficients of a high power of a known power series which is due to Euler [2]. The authors are grateful to H. Wilf for pointing out Lemma 2.5.

Lemma 2.5. *Let L be a field. Assume either (1) $q \in L$ and L has characteristic 0, or (2) q is an integer. Let $F = \sum_{n\geq 0} F_n t^n$ and $G = \sum_{n\geq 0} G_n t^n$ in $L[[t]]$ be formal power series with constant term 1 that satisfy $F = G^q$. Then*

$$\sum_{0\leq j\leq n} (n - (q+1)j)F_{n-j}G_j = 0. \tag{6}$$

Proof. Take logarithmic derivatives to get

$$\frac{F'}{F} = q\frac{G'}{G}.$$

Next, cross-multiply to get

$$F'Gt = qFG't,$$

and equate coefficients of t^n to get

$$\sum_{0\le j\le n-1}(n-j)F_{n-j}G_j=\sum_{1\le j\le n}qF_{n-j}jG_j.$$

There is no harm in including the term $j=n$ in the sum on the left and the term $j=0$ in the sum on the right, since these terms are zero. Now subtract to get equation (6). □

This proves Corollary 1.3.

3. Finite fields

Recall that R is a complete discrete valuation ring with finite residue field k, where $q=|k|$. Let n be a positive integer. Write $\bar{r}_n$ for the probability that a random monic degree n polynomial with coefficients in k splits completely. Then $\bar{r}_n$, as well as r_n, depends only on q and n. Our convention is that $r_0=\bar{r}_0=1$. We note some properties of r_n and $\bar{r}_n$.

1. We have $r_n\le\bar{r}_n$. For, if a monic polynomial f in $R[x]$ splits completely, then $\overline{f}$ splits completely in $k[x]$.

2. We have $\sum_{n\ge0}\bar{r}_nt^n=(1-t/q)^{-q}$. For, by the result sometimes called the stars and bars theorem, the number of monic degree n polynomials that split completely is $q^n\bar{r}_n=\binom{n+q-1}{q-1}$. Now use the binomial expansion $(1-t)^{-q}=\sum_{n\ge0}\binom{-q}{n}(-t)^n$, and the fact that $\binom{-q}{n}=(-1)^n\binom{q-1+n}{n}$.

3. We have $\lim_{q\to\infty}r_n=1/n!=\lim_{q\to\infty}\bar{r}_n$. The second equality follows from Observation 2. For f monic in $R[x]$, the probability that $\overline{f}$ has a repeated root is at most $1/q$. (Proof: the set S of monic polynomials f with constant and linear term in πR has measure $1/q^2$. Hence the union of the sets $\psi_r(S)$ over a set of lifts r of the elements of k has measure at most $1/q$.) But this tends to zero as $q\to\infty$. This proves the first equality.

4. Non-monic polynomials

In this section we drop the assumption that our polynomials are monic. We identify the set $P_n^{\mathrm{nm}}(R)$ of degree $\le n$ polynomials with coefficients in R with R^{n+1}. It gets the measure $\mu_n^{\mathrm{nm}}=\mu^{n+1}$. Let $S_n^{\mathrm{nm}}(R)$ be those polynomials $f(x)$ in $P_n^{\mathrm{nm}}(R)$ that can factored into n linear factors, i.e. can be written in the form

$$f=(b_1x-a_1)\cdots(b_nx-a_n),\tag{7}$$

with a_i, $b_i\in R$.

One shows that $S_n^{\mathrm{nm}}(R)$ is measurable as follows. Take a pair $(a_1,\dots,a_n)$ and $(b_1,\dots,b_n)$ of n-tuples and construct the polynomial f as in (7). The resulting continuous map from the compact set R^{2n} to $P_n^{\mathrm{nm}}(R)$ has image $S_n^{\mathrm{nm}}(R)$. Thus $S_n^{\mathrm{nm}}(R)$ is compact, hence closed.

The goal of this section is a formula for $r_n^{\mathrm{nm}}=\mu_n^{\mathrm{nm}}(S_n^{\mathrm{nm}}(R))$. This may be interpreted as the probability that a random, not necessarily monic, polynomial of degree $\le n$ splits completely into linear factors.

The first four values are:

$$\begin{aligned} r_1^{\mathrm{nm}} &= 1, \\ r_2^{\mathrm{nm}} &= 1/2, \\ r_3^{\mathrm{nm}} &= (q^2+1)^2(q-1)/6(q^5-1), \\ r_4^{\mathrm{nm}} &= h^{\mathrm{nm}}(q-1)^2/24(q^5-1)(q^9-1), \end{aligned}$$

where $h^{\mathrm{nm}} = q^{12}-q^{11}+4q^{10}+3q^8+4q^7-q^6+4q^5+3q^4+4q^2-q+1$. By convention, we set $r_0^{\mathrm{nm}} = 1$. We will need the following generalization of Lemma 2.3. Again it is an immediate consequence of Lemma 2.4.

Let g be a (not necessarily monic) polynomial in $k[x]$, with $\deg(g) \le n$. Write $P_{n,g}$ for the set of polynomials f in $R[x]$ of degree at most n whose reduction $\overline{f}$ is g. The space $P_{n,g}$ is a measurable subset of $P_n^{\mathrm{nm}}(R)$ and therefore inherits a topology and a measure. We renormalize this measure to give $P_{n,g}$ total measure 1.

Recall that if h in $k[x]$ is monic, then we defined P_h to be the set of monic polynomials f in $R[x]$ such that $\overline{f} = h$.

Lemma 4.1. *Let g and h in $k[x]$ be relatively prime polynomials of degree m and n, respectively, and assume that h is monic. Let $m' \ge m$. Then the multiplication map*

$$M : P_{m',g} \times P_h \to P_{m'+n,gh}$$

is a measure-preserving homeomorphism.

We will write k^* for the nonzero elements of the residue field k. Let $j = j(f)$ be the multiplicity of ∞ as a root of $\overline{f}$, by which we mean that exactly the first j leading coefficients of $\overline{f}$ are zero. Evidently, $0 \le j \le n+1$. Then if $j \le n$, by Hensel's lemma f factors uniquely as $f = f^{\mathrm{inf}} f_{\mathrm{inf}}$, where $\deg f_{\mathrm{inf}} \le j$, $\deg f^{\mathrm{inf}} = n-j$, f_{inf} reduces mod πR to a nonzero constant, and f^{inf} is monic.

Lemma 4.2. *If $0 \le i \le n$, then:*

(1) $\mathrm{pr}(f \text{ splits completely} | j(f) = i) =$

$$\mathrm{pr}(f_{\mathrm{inf}} \text{ splits completely} | \deg f_{\mathrm{inf}} \le i,\ \overline{f_{\mathrm{inf}}} \in k^*) \cdot$$
$$\mathrm{pr}(f^{\mathrm{inf}} \text{ splits completely} | \deg f^{\mathrm{inf}} = n-i,\ f^{\mathrm{inf}} \text{ is monic}).$$

(2) $\mathrm{pr}(f_{\mathrm{inf}} \text{ splits completely} | \deg f_{\mathrm{inf}} \le i,\ \overline{f_{\mathrm{inf}}} \in k^*) = q^{-\binom{i}{2}} r_i$.

(3) $\mathrm{pr}(f^{\mathrm{inf}} \text{ splits completely} | \deg f^{\mathrm{inf}} = n-i,\ f^{\mathrm{inf}} \text{ is monic}) = r_{n-i}$.

(4) *If $i = n+1$ (i.e., if $\overline{f} = 0$), then*

$$\mathrm{pr}(f \text{ splits completely} | j(f) = i) = r_n^{\mathrm{nm}}.$$

Proof. Part (1) of the lemma follows from Lemma 4.1.

After dividing $x^i f_{\mathrm{inf}}(1/x)$ by its leading coefficient, we get a random monic polynomial reducing to x^i. Applying part (2) of Lemma 2.2 then proves (2).

(3) is immediate by definition.

Conditioned on $\overline{f} = 0$, f/π is a random polynomial in $P_n^{\mathrm{nm}}(R)$. This proves (4). □

Now write $f = \sum_{0 \le i \le n} a_i x^{n-i}$. Then $j(f)$ is j if and only if $\overline{a_0} = \cdots = \overline{a_{j-1}} = 0$ and (if $j \le n$) $\overline{a_j} \ne 0$.

Thus, the probability that $j(f)$ equals j is $\frac{q-1}{q^{j+1}}$ if $0 \le j \le n$, and $\frac{1}{q^{n+1}}$ if $j = n+1$.

Thus, we have

$$r_n^{\mathrm{nm}} = \sum_{0\le j\le n} \frac{q-1}{q} r_{n-j} \frac{r_j}{q^{\binom{j+1}{2}}} + \frac{r_n^{\mathrm{nm}}}{q^{n+1}}. \tag{8}$$

This is easily seen to be equivalent to

$$F^{\mathrm{nm}} = \frac{q-1}{q} FG,$$

where

$$F^{\mathrm{nm}}(t) = \sum_{n\ge 0} (1-q^{-n-1}) r_n^{\mathrm{nm}} t^n, \text{ and} \tag{9}$$

$$G(t) = \sum_{n\ge 0} \frac{r_n}{q^{\binom{n+1}{2}}} t^n. \tag{10}$$

Here is another equivalent formulation:

Corollary 4.3. $F^{\mathrm{nm}} = \frac{q-1}{q} G^{q+1}$.

Using Euler's method, we get

$$\sum_{0\le j\le n} (n-(q+2)j)(1-q^{-(n-j)-1}) r_{n-j}^{\mathrm{nm}} \frac{r_j}{q^{\binom{j+1}{2}}} = 0. \tag{11}$$

In particular, r_n^{nm} is a rational function of q.

Here is a slightly different viewpoint, which leads to a slightly different formulation of the result.

Let $\mathbf{P}(k)$ be the set of lines through the origin in the plane k^2. Such a line is uniquely determined by its slope, which is an element of the set k or undefined. This identifies $\mathbf{P}(k)$ with $k \cup \{\infty\}$; hence, $\mathbf{P}(k)$ has cardinality $q+1$.

Assume for the moment that f splits completely, and that $\overline{f}$, its reduction mod πR, is nonzero. Write

$$\overline{f} = (\overline{b}_1 x - \overline{a}_1)\cdots(\overline{b}_n x - \overline{a}_n).$$

Then, for each i, $(\overline{a}_i, \overline{b}_i)$ in k^2 is not the origin, so the line through this point and the origin is in $\mathbf{P}(k)$. As i varies, we get a a set of n points in $\mathbf{P}(k)$, with multiplicities.

For $\ell \in \mathbf{P}(k)$, let d_ℓ be the number of $1 \le i \le n$ such that $(\overline{a}_i, \overline{b}_i)$ lies on ℓ. We have $n = \sum_{\ell\in\mathbf{P}(k)} d_\ell$.

For any function d' from $\mathbf{P}(k)$ to the set $\mathbf{N}$ of nonnegative integers, set $|d'| = \sum_{\ell\in\mathbf{P}(k)} d'_\ell$. Reasoning as in Lemmas 2.2 and 4.2 now yields

Theorem 4.4. *Let R be a complete discrete valuation ring with finite residue field k. Set $q = |k|$. For n a positive integer, write r_n^{nm} for the probability that a random degree $\le n$ polynomial over R splits into linear factors. Set $r_0^{\mathrm{nm}} = 1$. Then*

$$r_n^{\mathrm{nm}} = \frac{r_n^{\mathrm{nm}}}{q^{n+1}} + \frac{q-1}{q} \sum_{|d|=n} \prod_{i\in\mathbf{P}(k)} \frac{r_{d_i}}{q^{\binom{d_i+1}{2}}}.$$

5. Generating functions

In this section we regard q as a variable, and define the sequence of rational functions $(r_n) = (r_n(q))$ by $r_0 = r_1 = 1$, and by

$$F(t) = \sum_{n\geq 0} r_n t^n, \tag{12}$$

$$G(t) = \sum_{n\geq 0} \frac{r_n}{q^{\binom{n+1}{2}}} t^n, \qquad \text{and} \tag{13}$$

$$F(t) = \exp(q \log G(t)) = G(t)^q. \tag{14}$$

Thus, if we plug in a prime power for q we recover the r_n's with their previous meaning.

We note some properties of these rational functions.

Lemma 5.1.

(1) *The degree of the numerator of r_n is the degree of the denominator, and* $\lim_{q\to\infty} r_n(q) = 1/n!$.
(2) *r_n vanishes at 0 to order $\binom{n}{2}$.*
(3) *The only poles of r_n are at roots of unity.*
(4) $r_n(q) = r_n(1/q) q^{\binom{n}{2}}$.
(5) *If q is fixed, $|q| \geq 3$ and q is not an integer then, as $n \to \infty$,*

$$r_n(q) \sim (-z_0)^q c^q \Gamma(-q)^{-1} n^{-q-1} z_0^{-n}$$

for some complex numbers z_0, $c \neq 0$.

Proof of Properties (1)–(4). Property (1) follows from Observation 3 in Section 3.

For (4), take the defining equations (12), (13) and (14) of r_n, and substitute qt for t. We get

$$\sum_{n\geq 0} r_n(q) q^n t^n = \left(\sum_{n\geq 0} \frac{r_n(q)}{q^{\binom{n+1}{2}}} q^n t^n \right)^q,$$

which we can rewrite as

$$\sum_{n\geq 0} \frac{r_n(q)q^n}{q^{\binom{n+1}{2}}} \frac{1}{(1/q)^{\binom{n+1}{2}}} t^n = \left(\sum_{n\geq 0} \frac{r_n(q)q^n}{q^{\binom{n+1}{2}}} t^n \right)^q,$$

or as

$$\left(\sum_{n\geq 0} \frac{r_n(q)}{q^{\binom{n}{2}}} \frac{1}{(1/q)^{\binom{n+1}{2}}} t^n \right)^{1/q} = \sum_{n\geq 0} \frac{r_n(q)}{q^{\binom{n}{2}}} t^n.$$

But this is the defining equation for $r_n(1/q)$, with $r_n(q)/q^{\binom{n}{2}}$ substituted for $r_n(1/q)$. Also, for $n = 0$ and $n = 1$, $r_n(q)/q^{\binom{n}{2}} = 1$. Therefore we must have $r_n(1/q) = r_n(q)/q^{\binom{n}{2}}$.

For (3), define $\bar{s}_n = \bar{s}_n(q)$ by $\bar{s}_n(q) = r_n(q) q^{-\binom{n}{2}}$. We will prove by induction on n that the only poles of $\bar{s}_n$ are at roots of unity. Since $\bar{s}_0 = \bar{s}_1 = 1$, this is clear for $n < 2$. If $n \geq 2$, multiply Corollary 1.3 by q^n and rewrite it as

$$\sum_{0\leq j\leq n} (n - (q+1)j) \bar{s}_{n-j} \bar{s}_j q^{\binom{n-j+1}{2}} = 0.$$

Separating out the terms involving $\bar{s}_n$ gives

$$nq^{\binom{n+1}{2}}\bar{s}_n - nq\bar{s}_n = qt_n,$$

where t_n is a polynomial in q, $\bar{s}_1, \ldots, \bar{s}_{n-1}$. But then $\bar{s}_n = t_n/n(q^{\binom{n+1}{2}-1} - 1)$, so the result follows from the induction hypothesis.

Property (2) is an immediate consequence of (1) and (4). This finishes the proof of Properties (1)–(4). □

Property (5) will be proved in the remainder of the section. By property (3), it makes sense to plug in for q any complex number in the complement of the unit disk, $\mathcal{S} = \{q \,|\, |q| > 1\}$. For $q \in \mathcal{S}$ not a positive integer, we expect that the power series G has a complex zero z_0 (necessarily nonzero, as $r_0 = 1$) which is simple and is the unique zero in the disc $\{z \,|\, |z| \leq |z_0|\}$). In fact, this is the case if $|q| \geq 3$ by Theorem 5.4 below. Property (5) follows from this theorem by singularity analysis, with the constant $c = G'(z_0)$.

Our particular focus is on the asymptotics of $r_n(q)$, as $n \to \infty$. Surprisingly, the behavior in $\mathcal{S} \cap \mathbf{N}$ is radically different from the behavior in $\mathcal{S} \setminus \mathbf{N}$.

To explain this, for $q \in \mathcal{S}$, we note that the terms do not grow too fast.

Lemma 5.2. *For each $q \in \mathcal{S}$, there is some positive real $M = M(q)$ such that $|r_n(q)| \leq M^n$ for all n.*

Proof. Since $|q| > 1$, the series $\sum_{n\geq 1} |q|^{1-\binom{n+1}{2}}$ converges, so pick k such that $\sum_{n\geq k} |q|^{1-\binom{n+1}{2}} < 1$, and then pick $0 < \epsilon < 1$ such that

$$\epsilon \sum_{1\leq n<k} |q|^{1-\binom{n+1}{2}} < 1 - \sum_{n\geq k} |q|^{1-\binom{n+1}{2}}. \tag{15}$$

Finally, pick $M \geq 1$ large enough so that $|r_n| < \epsilon M^n$ for $n = 1, \ldots, k-1$. We will prove by induction that for all n, $|r_n| \leq M^n$. We may assume that $n \geq \max(k, 2)$. Now for $n \geq 2$, Corollary 1.3 tells us that

$$r_n = \frac{\sum_{1\leq j\leq n-1}(n-(q+1)j)r_{n-j}r_j q^{-\binom{j+1}{2}}}{n(q^{1-\binom{n+1}{2}} - 1)}.$$

However, $|n-(q+1)j| \leq n - j + |q|j = n + (|q|-1)j \leq n|q|$, so the absolute value of the numerator is bounded by

$$\sum_{1\leq j\leq n-1} n|r_{n-j}||r_j||q|^{1-\binom{j+1}{2}}$$

and by our assumption and the induction hypothesis this is less than

$$nM^n(\epsilon \sum_{1\leq j<k} |q|^{1-\binom{j+1}{2}} + \sum_{k\leq j\leq n-1} |q|^{1-\binom{j+1}{2}}).$$

Now the absolute value of the denominator is no smaller than $n(1 - |q|^{1-\binom{n+1}{2}})$, so to have $|r_n| \leq M^n$ it will do to have

$$\epsilon \sum_{1\leq j<k} |q|^{1-\binom{j+1}{2}} + \sum_{k\leq j\leq n-1} |q|^{1-\binom{j+1}{2}} \leq 1 - |q|^{1-\binom{n+1}{2}},$$

which follows immediately from (15). This completes the proof. □

From this it follows that the series G has an infinite radius of convergence. If $q > 1$ is an integer, then $F = G^q$ implies that F also has an infinite radius of convergence, whence its coefficients are small: $r_n(q) = O(\epsilon^n)$ for all $\epsilon > 0$. (In fact, much more is true, as we shall see in the next section.) For q nonintegral, on the other hand, property (5) says that this will not in general be true.

Lemma 5.3. *If $|q| \geq 2$, $|r_n(q)| \leq 1$ for all n.*

Proof. We proceed by induction on n. For $n = 0$ and $n = 1$ the result is trivial. So let $n \geq 2$. Corollary 1.3 then gives the bound

$$|r_n| \leq \frac{\sum_{1\leq j\leq n-1}(n-j+|q|j)|r_{n-j}||r_j||q|^{-\binom{j+1}{2}}}{n(1-|q|^{1-\binom{n+1}{2}})},$$

and by the induction hypothesis, this is no bigger than

$$\frac{\sum_{1\leq j\leq n-1}(n-j+|q|j)|q|^{-\binom{j+1}{2}}}{n(1-|q|^{1-\binom{n+1}{2}})},$$

so it will do to show that

$$\sum_{1\leq j\leq n}(n-j+|q|j)|q|^{-\binom{j+1}{2}} \leq n.$$

Since the left-hand side of this inequality is nonincreasing in $|q|$, it will do to show that

$$\sum_{1\leq j\leq n}(n+j)2^{-\binom{j+1}{2}} \leq n.$$

For $n = 2$ this can be directly verified. Otherwise, $n \geq 3$. Now observe that

$$\sum_{1\leq j\leq n} n2^{-\binom{j+1}{2}} \leq \sum_{j\geq 1} n2^{-\binom{j+1}{2}} \leq \sum_{j\geq 1} n2^{-(2j-1)} = \frac{2n}{3}$$

and

$$\sum_{1\leq j\leq n} j2^{-\binom{j+1}{2}} \leq \sum_{j\geq 1} j2^{-(2j-1)} = \frac{8}{9} \leq 1 \leq \frac{n}{3}.$$

This completes the proof. □

Theorem 5.4. *If $q = 2$ or $|q| \geq 3$, $G(z)$ has a unique zero in the disc $D = \{z \mid |z| < |q|+1\}$, which is simple.*

Proof. Set $G_0(z) = r_0 + r_1 z/q + r_2 z^2/q^3 = 1 + z/q + z^2/(2q^2(q+1))$. Then G_0 has roots at $-(q+q^2) \pm q^2\sqrt{1-q^{-2}}$. By looking at the power series for $\sqrt{1+x}$, we can write $\sqrt{1-q^{-2}} = 1 - q^{-2}/2 + \epsilon$, where $|\epsilon| < |q|^{-4}/5$. Plugging in this estimate yields roots $r_1 = -q - 1/2 + q^2\epsilon$ and $r_2 = -2q^2 - q + 1/2 - q^2\epsilon$, and since

$$|r_1| \leq |q| + \frac{1}{2} + \frac{|q|^{-2}}{5} \leq |q| + \frac{11}{20} < |q| + 1$$

and

$$|r_2| > 2|q|^2 - |q| - \frac{1}{2} - \frac{|q|^{-2}}{5} \geq 2|q|^2 - |q| - \frac{11}{20} > |q| + 1,$$

exactly one of these roots is in D. By Rouché's Theorem, it will do to show that $|G(z) - G_0(z)| < |G_0(z)|$ on C, where $C = \{z \mid |z| = |q| + 1\}$. Let $z \in C$. We then have

$$\begin{aligned} |(q+1)G_0(z)| &\geq \frac{1}{2}|q^{-2}|(|z| - |r_1|)(|r_2| - |z|) \\ &\geq \frac{1}{2}|q|^{-2}\frac{9}{20}(2|q|^2 - 2|q| - \frac{31}{20}) \\ &\geq \frac{9}{20}(1 - |q|^{-1} - \frac{31|q|^{-2}}{40}) \\ &\geq \frac{441}{3200}. \end{aligned} \tag{16}$$

Now,

$$|(q+1)r_3\frac{z^3}{q^6}| = \frac{1}{6}(1 + \frac{1}{|q|})^3\frac{|q^2 - q + 1||q - 1|}{|q^5 - 1|}. \tag{17}$$

If $q = 2$, the right-hand side of (17) is $27/496$, and if $|q| \geq 3$, it is no more than

$$\frac{1}{6}(1 + \frac{1}{|q|})^3|q|^{-2}\frac{(1 + |q|^{-1} + |q|^{-2})(1 + |q|^{-1})}{1 - |q|^{-5}} \leq \frac{1}{6}(\frac{4}{3})^3 3^{-2}\frac{(13/9)(4/3)}{1 - 3^{-5}} = \frac{832}{9801}.$$

In any case, it is no more than $9/100$. Again, we have

$$|(q+1)r_4\frac{z^4}{q^{10}}| = \frac{1}{24}(1 + \frac{1}{|q|})^4\frac{|h(q)||q - 1|^2}{|q + 1||q^5 - 1||q^9 - 1|}. \tag{18}$$

If $q = 2$, the right-hand side of (18) is $1161/2027648$. If $|q| \geq 3$, $|q^6 h(q)| \leq 3|(q^5 - 1)(q^9 - 1)|$, so it is no more than

$$\frac{1}{24}(1 + \frac{1}{|q|})^4|q|^{-5}\frac{3(1 + |q|^{-1})^2}{1 - |q|^{-1}} \leq \frac{1}{24}(\frac{4}{3})^4 3^{-5}\frac{3(1 + 3^{-1})^2}{1 - 3^{-1}} = \frac{256}{59049}.$$

In any case, it is no more than $1/200$. Finally, using Lemma 5.3, we have

$$\begin{aligned} \left|\sum_{n\geq 5}\frac{(q+1)r_n}{q^{n(n+1)/2}}z^n\right| &\leq |1 + q^{-1}|\sum_{n\geq 5}|\frac{z}{q}|^n|q|^{-(n(n-1)/2-1)} \\ &\leq (1 + |q|^{-1})\sum_{n\geq 5}(1 + |q|^{-1})^n|q|^{-(n(n-1)/2-1)} \\ &\leq \frac{3}{2}\sum_{n\geq 5}(\frac{3}{2})^n 2^{-(n(n-1)/2-1)} \\ &\leq (\frac{3}{2})^6 2^{-9}\frac{1}{1 - \frac{3}{2}2^{-5}} \\ &= \frac{729}{31232} \leq \frac{1}{40}. \end{aligned}$$

Adding up these inequalities tells us that

$$|(q+1)(G_0(z) - G(z))| = \left|\sum_{n\geq 3}\frac{(q+1)r_n}{q^{n(n+1)/2}}z^n\right| \leq \frac{9}{100} + \frac{1}{200} + \frac{1}{40} = \frac{3}{25}.$$

Together with (16), this proves the result. $\square$

As remarked above, property (5) follows from Theorem 5.4.

6. Open questions

We mention some open questions for further research.

1. Zeroes of the numerator. Let N_n be the numerator of r_n. Is it true that N_n has no zeroes on the negative real axis?

2. Zeroes of G. Let $q \geq 2$ be an integer. Is it true that the zeroes of the function G are all real and negative?

We remark that the zeroes $|z_0| \leq |z_1| \leq \cdots$ of G completely determine G as follows. Because $G(t)$ is an entire function of order zero, i.e., for all $\alpha > 0$, $|G(z)| = o(e^{|z|^\alpha})$ as $|z| \to \infty$, it is equal up to a constant factor to the Weierstrass product $\prod_{i\geq 0}(1 - \frac{t}{z_i})$. But then $G(t)$ equals this product since $G(0) = 1$.

3. Asymptotics. Prove the more accurate asymptotic expansion (44) for $\log_q r_n(q)$, for $q > 2$ an integer.

7. Asymptotics

In this section $q \geq 2$ is an integer. If $q = 2$, we give the first three terms in the asymptotic expansion of $\log_2 r_n$. If $q > 2$, we give the first two terms in the asymptotic expansion of $\log_q r_n$ and a guess for its third term.

Theorem 7.1. *If $q = 2$, then there is a constant K_2 and a continuous function $\bar{W}_2$ with period 1 such that*

$$\left|\log_2 r_n + \frac{n^2}{2} + \frac{1}{2} n \log_2 n - (\frac{1}{2} + \bar{W}_2(\log_2 n))n\right| \leq K_2, \tag{19}$$

for all $n \geq 1$. (Figure 2 shows a graph of $\bar{W}_2$.) If $q \geq 2$ is an integer, then there is a constant C_q such that

$$\left|\log_q r_n + \frac{n^2}{2(q-1)} + \frac{1}{2} n \log_q n\right| \leq C_q n, \tag{20}$$

for all $n \geq 1$, and we conjecture (Conjecture 9.4) that there is a constant K_q and a continuous function $\bar{W}_q$ with period 1 such that

$$\left|\log_q r_n + \frac{n^2}{2(q-1)} + \frac{1}{2} n \log_q n - (\frac{1}{2} + \bar{W}_q(\log_q n))n\right| \leq K_q,$$

for all $n \geq 1$.

The idea of the proof of (20) is this. In the proof of the exact formula for r_n, we looked at a certain infinite tree whose branching was given by the factorization of a polynomial mod the maximal ideal πR of R, mod $\pi^2 R$, etc. In a modification of this method, we use finite trees instead. Lemma 7.6 expresses $r_n/q^{\binom{n+1}{2}}$ as the sum of a certain function H over labelled q-trees (defined below) with n leaves. By Lemma 7.3, we will see that for any fixed q, the logarithm of the number of labelled q-trees is $O(n)$. For example, consider the case $q = 2$. In this case our labelled q-trees are the same as binary trees. The number of binary trees with n leaves is given by the Catalan number $\frac{1}{n}\binom{2n-2}{n-1}$. By Stirling's formula, this has logarithm $O(n)$.

If we set M to be the maximum value of $\log_q H$, we then have

$$M \leq -\binom{n+1}{2} + \log_q r_n \leq M + O(n).$$

Thus, to prove (20), it will be enough to estimate the maximum value of $\log_q H$ to within $O(n)$. We will see by Lemma 7.11 that H is maximized at a tree we will call the *well-balanced q-tree*, and in Subsection 7.4 we will calculate H for the well-balanced q-tree.

Although (19) is a refinement of (20), its proof is different, and relies on a direct use of (1), together with an estimate of how much the largest term appearing in (1) contributes to (1).

7.1. **q-trees.** A rooted tree is a connected acyclic graph with a distinguished vertex (the root). We can direct the edges of a rooted tree in a unique way, by directing them away from the root. The root has in-degree 0; all other vertices have in-degree 1. The edges emanating from a vertex go to distinct vertices, called the *children* of v. A vertex with no children is a *leaf* vertex.

A *q-tree* is a rooted tree in which the number of out-edges from each vertex is at most q, and is not 1 (i.e. there is branching at each vertex that is not a leaf.)

For v a vertex of the q-tree T, we write T_v for the full subtree whose vertices are v and all its descendants. This is a q-tree.

A *labelled q-tree* is a q-tree together with a labelling of its edges with elements of the set $S = \{0, 1, \dots, q-1\}$ so that the out-edges emanating from each vertex have distinct labels.

Example 7.2. We list the q-trees with at most 3 leaves.

(i) The tree τ_1 with one vertex and no edges is the unique q-tree with one leaf. In all other q-trees, a vertex is a leaf if and only if it has total degree 1.
(ii) The tree τ_2 with three vertices, a root with two leaf children, is the unique q-tree with two leaves. There are $\binom{q}{2}$ ways of labelling this q-tree.
(iii) There are two q-trees with three leaves (one if $q = 2$). One, call it τ_{3a}, has a root with three leaf children (if $q > 2$). The other, τ_{3b}, has a root with two children v and w, with v a leaf and $T_w = \tau_2$. There are $\binom{q}{3}$ ways of labelling τ_{3a}, and $q(q-1)\binom{q}{2}$ ways of labelling τ_{3b}.

Lemma 7.3. *The number of labelled q-trees with l leaves is at most* $(2q+1)^{5l-3}$.

Proof. Let Σ be the set of $2q+1$ symbols: the parentheses $(_i$ and $)_i$ for $0 \le i \le q-1$, plus a dot. We construct for every labelled q-tree a distinct string of at most $5l-4$ of these symbols, thus constructing an injective map from the set of labelled q-trees with l leaves to $\cup_{0 \le i \le 5l-4} \Sigma^i$.

We proceed by induction on l. For $l = 1$, the tree τ_1 corresponds to the dot.

Assume $l \ge 2$, and let T be a labelled q-tree with l leaves. Let the root of T have j children. For each child v of the root of T, bracket the sequence corresponding to the subtree T_v with the symbols $(_i$ and $)_i$, where $i \in S$ is the label of the edge from the root to the vertex v. By the induction hypothesis, this uses at most $5l - 4j + 2j = 5l - 2j \le 5l - 4$ symbols.

The number of labelled q-trees with l leaves is therefore at most

$$\sum_{0 \le i \le 5l-4} (2q+1)^i = (2q)^{-1}((2q+1)^{5l-3} - 1) \le (2q+1)^{5l-3}.$$

$\square$

7.2. **A q-tree recursion.** Equating coefficients of t^n in the defining equations (12), (13) and (14) of r_n gives us

$$r_n = \sum_{|b|=n} \prod_{i=0}^{q-1} \frac{r_{b_i}}{q^{\binom{b_i+1}{2}}}. \tag{21}$$

Group the r_n terms to get

$$r_n = \frac{q}{q^{\binom{n+1}{2}}} r_n + {\sum_{|b|=n}}' \prod_{i=0}^{q-1} \frac{r_{b_i}}{q^{\binom{b_i+1}{2}}},$$

where $\sum'$ is the sum over b such that $b_i > 0$ for at least two values of i.

Setting $s_n = r_n / q^{\binom{n+1}{2}}$, we have

$$q^{\binom{n+1}{2}} s_n = q s_n + {\sum_{|b|=n}}' \prod_{i=0}^{q-1} s_{b_i}.$$

For $n \geq 2$, set $\beta_n = 1/(q^{\binom{n+1}{2}} - q)$. We have proved the following.

Lemma 7.4. *Assume $n \geq 2$. Then,*

$$s_n = \beta_n {\sum_{|b|=n}}' \prod_{i=0}^{q-1} s_{b_i}. \tag{22}$$

Remark 7.5. We can decompose a q-tree into its root, plus a subtree T_v for each child v of the root of T. This decomposition underlies Lemma 7.3, and will allow us to interpret Lemma 7.4 as a recursion on labelled q-trees.

Write $\ell(v)$ for the number of leaves of the tree T_v. As an easy application of the decomposition in Remark 7.5, we have the following. Let T be a q-tree with more than one vertex. Then the root of T is not a leaf, and the number of leaves of T equals $\sum \ell(v)$, where the sum is over the set of children v of the root of the tree T.

We can now interpret the recursion for s_n in terms of labelled q-trees. Set $\beta_1 = s_1 = 1/q$.

Lemma 7.6. *We have*

$$s_n = \sum \prod_{v \in T} \beta_{\ell(v)}, \tag{23}$$

where the sum is over labelled q-trees T with n leaves, and the product is over all vertices v of the tree T.

Proof. For $n = 1$, this follows by our choice of β_1. The general case follows by Remark 7.5 and Lemma 7.4. $\square$

We shall write $H(T)$ for $\prod_{v \in T} \beta_{\ell(v)}$.

Example 7.7. We recalculate r_1, r_2 and r_3 using Lemma 7.6.

(i) We have $s_1 = 1/q$, so that $r_1 = 1$.
(ii) We have $H(\tau_2) = \beta_1^2 \beta_2$. There are $\binom{q}{2}$ ways of labelling the edges of τ_2. Hence $r_2 = q^3 s_2 = q^3 \binom{q}{2} \beta_1^2 \beta_2 = q/2(q+1)$.

(iii) We have $H(\tau_{3a}) = \beta_3\beta_1^3$. There are $\binom{q}{3}$ ways of labelling the edges of τ_{3a} (even if $q = 2$!). The contribution to r_3 from τ_{3a} is thus $q^6\binom{q}{3}\beta_3\beta_1^3$. We have $H(\tau_{3b}) = \beta_3\beta_2\beta_1^3$. There are $q(q-1)\binom{q}{2}$ ways of labelling the edges of τ_{3b}. Hence, the contribution to r_3 from τ_{3b} is $q^6q(q-1)\binom{q}{2}\beta_3\beta_2\beta_1^3$. Add, and this agrees with the formula in Section 1.

7.3. **The well-balanced q-tree.** In Subsection 7.2 we defined $\beta_n = 1/(q^{\binom{n+1}{2}} - q)$ for $n \geq 2$ and $\beta_1 = 1/q$.

Lemma 7.8. *We have*

(i): $\log_q \beta_n = -\binom{n+1}{2} + o(1)$ *as* $n \to \infty$.
(ii): *The sequence* $(\beta_n/\beta_{n-1})_{n=2}^{\infty}$ *is monotone decreasing.*

Proof. Part (i) is clear since $\beta_n q^{\binom{n+1}{2}}$ tends to 1 as $n \to \infty$.

In order to prove (ii) it suffices to show $\frac{1}{\beta_n^2} \leq \frac{1}{\beta_{n-1}\beta_{n+1}}$ for all $n \geq 2$.

First consider the case $n = 2$. We have $\frac{1}{\beta_2^2} = (q^3 - q)^2 \leq (q^3 - q)(q^3 + q) = q^6 - q^2 \leq q(q^6 - q)$, and this last quantity is equal to $1/\beta_1\beta_3$.

Next suppose $n \geq 3$. Then $\frac{1}{\beta_n^2} = (q^{\binom{n+1}{2}} - q)^2$, and we have

$$\begin{aligned}
(q^{\binom{n+1}{2}} - q)^2 &\leq (q^{\binom{n+1}{2}} - q)(q^{\binom{n+1}{2}} + q) \\
&= q^{n^2+n} - q^2 \\
&\leq (q^{\binom{n}{2}-1} + q)(q^{\binom{n+2}{2}} - q) = AB, \qquad \text{say.}
\end{aligned}$$

The last inequality follows since $\binom{n}{2} - 1$ is less than $\binom{n+2}{2}$ and their sum is $n^2 + n$.

Since $2 \leq q$, we have $2 \leq q(q-1)$, which implies that $2q \leq q^j(q-1)$ for any $j \geq 2$. Rearranging the terms gives $q^j + q \leq q^{j+1} - q$, which implies (taking $j = \binom{n}{2} - 1 \geq \binom{3}{2} - 1 = 2$) that $A \leq 1/\beta_{n-1}$. As $B = 1/\beta_{n+1}$, the lemma is proved. $\square$

Let $n \geq 1$ be an integer. We are interested in q-tuples of nonnegative integers that sum to n, and such that each pair differs by at most one. We remark that such a q-tuple i exists: write $n = qx + y$ with $0 \leq y \leq q$ (we allow either $y = 0$ or $y = q$); now take $i_1 = \cdots = i_y = x + 1$ and $i_{y+1} = \cdots = i_q = x$. Moreover, if such a q-tuple contained a value less than x (resp. larger than $x + 1$), all values would be at most x (resp. at least x), and the sum would be too small (resp. too large). Thus all values are x or $x + 1$, and the number of each must be $q - y$ (resp. y). Hence the q-tuple is unique up to order.

Lemma 7.9. *Let $n \geq 1$ be an integer. There is a unique q-tree $T = T(n)$ with n leaves such that for every vertex v of T:*

(1) *If $\ell(v) < q$ then all children of v are leaves.*
(2) *If $\ell(v) \geq q$ then v has q children and, for any two children w and w' of v, $\ell(w)$ and $\ell(w')$ differ by at most 1.*

Proof. If $n = 1$, then T is the unique q-tree with one (leaf) vertex. If $1 < n < q$, then T is the unique q-tree with $n+1$ vertices that consists of a root with n children all of which are leaves.

For $n \geq q$, we define $T = T(n)$ by induction on n. Write $n = qx + y$ with $0 \leq y \leq q - 1$ as in the remarks preceding this lemma. Applying property 2 above at the root of T, we see that if T exists, its root must have children $v_1, \ldots, v_q$ such

that $\ell(v_1) = \cdots = \ell(v_y) = x+1$ and $\ell(v_{y+1}) = \cdots = \ell(v_q) = x$. It now follows from the induction hypothesis that each T_{v_i} must equal $T(\ell(v_i))$. This gives a unique candidate for T, namely the q-tree whose root has precisely q children, and such that $T_w = T(x+1)$ for y of these children w and $T_w = T(x)$ for the remaining $q-y$ children w. It is easy to see that this does indeed satisfy properties 1 and 2. $\square$

We call a q-tree *well-balanced* if it satisfies the two conditions of the lemma. For T the well balanced q-tree with $n \geq 1$ leaves, write $\nu_n = H(T) = \prod_{v\in T} \beta_{\ell(v)}$. Set $\nu_0 = 1$.

Lemma 7.10.

(i) $\nu_1 = 1/q$.

(ii) *If $n \geq 2$ and we write $n = qx + y$, $0 \leq y \leq q$, then*

$$\nu_n = \beta_n \nu_x^{q-y} \nu_{x+1}^y.$$

(iii) *The sequence $(\nu_i/\nu_{i-1})_{i=1}^\infty$ is monotone decreasing.*

Proof. (i) is obvious, and (ii) follows immediately from the defining properties of the well-balanced q-tree and the remarks preceding Lemma 7.9. We now prove that for all n, the sequence $(\nu_i/\nu_{i-1})_{i=1}^n$ is monotone decreasing; this will prove (iii). We induce on n. If $n \leq 3$, the result follows from Example 7.7, so assume that $n \geq 4$. We need to show that $\nu_{n-1}/\nu_{n-2} \geq \nu_n/\nu_{n-1}$.

For any $2 \leq m \leq n-1$, write $m = qx + y$ with $0 \leq y \leq q-1$ (note that we do not allow $y = q$). By (ii), $\nu_m = \beta_m \nu_x^{q-y} \nu_{x+1}^y$.

Next we calculate ν_{m+1}. We have $m + 1 = qx + (y + 1)$, with $y + 1 \leq q$, so (ii) gives $\nu_{m+1} = \beta_{m+1} \nu_x^{q-y-1} \nu_{x+1}^{y+1}$, whence $\nu_{m+1}/\nu_m = \frac{\beta_{m+1}}{\beta_m} \frac{\nu_{x+1}}{\nu_x}$.

We apply this to $m = n-2$ and $m = n-1$. The desired result now follows from part (ii) of Lemma 7.8, after possibly using the induction hypothesis. $\square$

Let γ_n be the largest tree contribution $H(T) = \prod_{v\in T} \beta_{\ell(v)}$ among q-trees T with n leaves. Note that here the labelling of the trees is irrelevant. We set $\gamma_0 = 1$. The following lemma says that the well-balanced q-tree always yields a largest tree contribution.

Lemma 7.11. *For all $n \geq 0$, $\gamma_n = \nu_n$.*

Proof. We proceed by induction on n. For $n = 0$ and $n = 1$, the result is obvious, so let $n > 1$. It will do to prove that $\gamma_n \leq \nu_n$.

The contribution from a tree with n leaves is at most $\beta_n \gamma_{i_1} \cdots \gamma_{i_q}$, where i_1, …, i_q are nonnegative integers that sum to n, at least two of which are positive. Since at least two i_z's are positive, we have $i_1, \ldots, i_q \leq n-1$, so by the induction hypothesis, $\gamma_{i_z} = \nu_{i_z}$ for all z. It therefore follows that the contribution is no more than

$$\beta_n \nu_{i_1} \cdots \nu_{i_q}. \tag{24}$$

Now by part (iii) of Lemma 7.10,

$$\nu_{i-1}\nu_j \leq \nu_i \nu_{j-1} \text{ if } 1 \leq i \leq j. \tag{25}$$

Let $n = qx + y$ with $0 \leq y \leq q$. It follows from (25) that (24) is largest when $|i_z - i_{z'}| \leq 1$ for all $1 \leq z, z' \leq q$, when, by the remarks preceding Lemma 7.9, it equals $\beta_n \nu_x^{q-y} \nu_{x+1}^y$. By Lemma 7.10, this equals ν_n, so we are done. $\square$

Lemma 7.12. *Let T be the well-balanced q-tree with n leaves, let $k \geq 0$, and let $n = q^k x + y$, $0 \leq y < q^k$. Then:*

1. *If $n < q^{k-1}$, there are no vertices at distance k from the root of T.*
2. *If $q^{k-1} \leq n < 2q^{k-1}$, there are $2(n - q^{k-1})$ vertices at distance k from the root of T, each of which is a leaf.*
3. *If $2q^{k-1} \leq n < q^k$, there are n vertices at distance k from the root of T, each of which is a leaf.*
4. *If $n \geq q^k$, there are q^k vertices v' at distance k from the root of T, y with $\ell(v') = x + 1$ and $q^k - y$ with $\ell(v') = x$.*

Proof. We prove the lemma by induction on k. If $k = 0$, these results are clear. Otherwise:

To prove the first claim, suppose that $n < q^{k-1}$. Then by the induction hypothesis, all the vertices at distance $k - 1$ from the root of T are leaves, and therefore there are no vertices at distance k from the root, as desired.

If $q^{k-1} \leq n < 2q^{k-1}$, then by the induction hypothesis, there are q^{k-1} vertices v' at distance $k-1$ from the root, $2q^{k-1} - n$ of which are leaves and $n - q^{k-1}$ of which satisfy $\ell(v') = 2$, i.e. have 2 children, both leaves; this gives a total of $2(n - q^{k-1})$ vertices at distance k from the root, all leaves, which is the second claim.

If $2q^{k-1} \leq n < q^k$, then write $n = q^{k-1}x' + y'$, $2 \leq x' \leq q - 1$, $0 \leq y' < q^{k-1}$. By the induction hypothesis, there are q^{k-1} vertices v' at distance $k - 1$ from the root, y' of which satisfy $\ell(v') = x' + 1$, and $q^{k-1} - y'$ of which satisfy $\ell(v') = x'$. By the properties of the well-balanced q-tree, the vertices v' for which $\ell(v') = x'$ must have x' children, all leaves. Similarly, the vertices v' for which $\ell(v') = x' + 1$ must have $x' + 1$ children, all leaves. This gives n vertices at distance k from the root, all leaves. This proves the third claim.

Finally, if $n \geq q^k$, write $y = y_0 q^{k-1} + y_1$, $0 \leq y_0 < q$, $0 \leq y_1 < q^{k-1}$. We then have $n = q^{k-1}(qx + y_0) + y_1$. By the induction hypothesis, therefore, there are q^{k-1} vertices v' at distance $k-1$ from the root, $q^{k-1} - y_1$ of which satisfy $\ell(v') = qx + y_0$. Since $x \geq 1$, $qx + y_0 \geq q$, so each of these vertices will, by the properties of the well-balanced q-tree, have y_0 children v'' satisfying $\ell(v'') = x + 1$ and $q - y_0$ children v'' satisfying $\ell(v'') = x$. Similarly, y_1 of the vertices at distance $k - 1$ from the root will satisfy $\ell(v') = qx + y_0 + 1$ and will have $y_0 + 1$ children v'' satisfying $\ell(v'') = x + 1$ and $q - y_0 - 1$ children v'' satisfying $\ell(v'') = x$. This gives a total of $y_0(q^{k-1} - y_1) + (y_0 + 1)y_1 = q^{k-1}y_0 + y_1 = y$ vertices v'' at distance k from the root with $\ell(v'') = x + 1$ and $(q - y_0)(q^{k-1} - y_1) + (q - y_0 - 1)y_1 = (q - y_0)q^{k-1} - y_1 = q^k - y$ vertices v'' at distance k from the root with $\ell(v'') = x$, as desired. This proves the final claim. □

7.4. **The largest tree contribution.** We wish to estimate the contribution from the well-balanced q-tree. By the previous subsection, the well-balanced q-tree gives the largest contribution to the sum (23).

The bound (20) in Theorem 7.1 will follow from:

Lemma 7.13. $\log_q \nu_n$*, the logarithm of the contribution to* (23) *from the well-balanced q-tree, is*

$$-\frac{n^2}{2(1 - q^{-1})} - \frac{n \log_q n}{2} + O(n). \tag{26}$$

Proof. We first treat the contribution η to ν_n from vertices at distance more than $\log_q n$ from the root. By Lemma 7.12, there are $O(n)$ such vertices, and each contributes a factor of β_1 to $\nu_n = \prod_v \beta_{\ell(v)}$. Thus $\log_q \eta = O(n)$.

Let k be an integer such that $1 \leq q^k \leq n$. Write ω_k for the contribution to ν_n from the vertices at distance k from the root.

Write $n = q^k x + y$, with $0 \leq y \leq q^k - 1$. By Lemma 7.12, we have $\omega_k = \beta_x^{q^k - y} \beta_{x+1}^y$, so that

$$\begin{aligned} \log_q \omega_k &= (q^k - y) \log_q \beta_x + y \log_q \beta_{x+1} \\ &= -(q^k - y)\binom{x+1}{2} - y\binom{x+2}{2} + O(q^k) \end{aligned}$$

by part (i) of Lemma 7.8. So, up to terms of order q^k, we have

$$\begin{aligned} \log_q \omega_k &= -\frac{x+1}{2}[(q^k - y)x + y(x+2)] \\ &= -\frac{x+1}{2}[n+y] \\ &= -\frac{1}{2}(n + y + \frac{n-y}{q^k}(n+y)) \\ &= -\frac{1}{2}(\frac{n^2}{q^k} + n). \end{aligned}$$

Now sum over k such that $1 \leq q^k \leq n$. Note that $\sum_k q^k$ is $O(n)$.
We get

$$\log_q \nu_n = \sum_{1 \leq q^k \leq n} \log_q \omega_k + \log_q \eta = -\frac{n^2}{2(1-q^{-1})} - \frac{n \log_q n}{2} + O(n). \tag{27}$$

The lemma is proved. $\square$

The bound (20) in Theorem 7.1 is now proved.

8. A recursion

We consider a recursion of the form

$$\begin{aligned} \Omega_n &= (q-y)\Omega_x + y\Omega_{x+1} + \omega_n, \\ &\qquad \text{where } n \geq 2,\ n = qx + y,\ 0 \leq y \leq q, \\ \Omega_0 &= 0. \end{aligned} \tag{28}$$

We now show how to solve this recursion if the sequence (ω_n) does not grow too rapidly with n.

Theorem 8.1. *Suppose $\omega_n = O(n^\kappa)$ for $0 \leq \kappa < 1$. The solution to (28) has the form*

$$\Omega_n = W(\log_q n)n + O(n^\kappa),$$

where W is a continuous function with period 1.

Proof. Let $n \geq 1$. It is clear from (28) that, if we set $\omega_1 = \Omega_1$, and let T be the well-balanced q-tree with n leaves, then Ω_n is the sum over all vertices v of T of $\omega_{\ell(v)}$. Write $\lfloor z \rfloor$ for the largest integer which is no more than z and $\{z\}$ for $z - \lfloor z \rfloor$. Now define $X(m)$ by

$$\begin{array}{lcll} X(m) & = & 0, & m < q^{-1}; \\ X(m) & = & 2\omega_1(m - q^{-1}), & q^{-1} \leq m < 2q^{-1}; \\ X(m) & = & \omega_1 m, & 2q^{-1} \leq m < 1; \\ X(m) & = & \omega_{1+\lfloor m \rfloor}\{m\} + \omega_{\lfloor m \rfloor}(1 - \{m\}), & 1 \leq m. \end{array}$$

Observe that X is continuous and that $X(m) = O(m^\kappa)$. Now Lemma 7.12 implies that the contribution to Ω_n from vertices at distance k from the root of T is $q^k X(n/q^k)$, so

$$\Omega_n = \sum_{k \geq 0} q^k X(\frac{n}{q^k}).$$

Note that the summand is zero for $k > Z = \lfloor \log_q n \rfloor + 1$. Now

$$\frac{\Omega_n}{n} = \sum_{0 \leq k \leq Z} \frac{X(n/q^k)}{n/q^k} \tag{29}$$

$$= \sum_{-Z \leq k \leq 0} \frac{X(nq^k)}{nq^k} \tag{30}$$

$$= \sum_{k \geq -Z} \frac{X(nq^k)}{nq^k} + O(n^{\kappa-1}), \qquad \text{as } X(m)/m = O(m^{\kappa-1}).$$

Set

$$W(m) = \sum_{k \geq -1-\lfloor m \rfloor} \frac{X(q^{m+k})}{q^{m+k}}, \tag{31}$$

or equivalently,

$$W(m) = \zeta(\{m\}), \qquad \zeta(x) = \sum_{k \geq -1} \frac{X(q^{x+k})}{q^{x+k}}. \tag{32}$$

It is now clear from (29) and (31) that $\Omega_n = W(\log_q n)n + O(n^\kappa)$. However, from (32), $X(m)/m = O(m^{\kappa-1})$, and the continuity of X, we see that on $[0,1]$, ζ is a sum of continuous functions which are uniformly bounded by a convergent series. Therefore ζ is continuous on $[0,1]$. The continuity of W now follows from the continuity of ζ on $[0,1]$ and the fact that, since $X(q^{-1}) = 0$, $\zeta(0) = \zeta(1)$. This concludes the proof. □

9. The third term

We can rewrite part (ii) of Lemma 7.10 as

$$\log_q \nu_n = \log_q \beta_n + (q-y)\log_q \nu_x + y \log_q \nu_{x+1}, \qquad \text{where } n \geq 2,\ n = qx + y,\ 0 \leq y \leq q. \tag{33}$$

Empirically, we have observed that a similar recursion appears to hold for $\log_q s_n$. This is because the main contribution in (22) comes when all the b_i's differ by at most 1. For $q = 2$, we can prove this.

Lemma 9.1. *If $q = 2$ and $n \geq 2$, then $s_{n+1}s_{n-1} \leq \frac{1}{2}s_n^2$.*

Proof. Set

$$R_m = \frac{s_{m+1}/s_m}{s_m/s_{m-1}} = \frac{s_{m+1}s_{m-1}}{s_m^2} \qquad (m \geq 2). \tag{34}$$

Then

$$\frac{s_{m+1}/s_m}{s_{m-(i-1)}/s_{m-i}} = R_m \cdots R_{m-(i-1)} \qquad (m \geq i+1,\ i \geq 0),$$

so

$$\frac{s_{m+j}s_{m-i}}{s_m s_{m+j-i}} = \frac{s_{m+j}/s_m}{s_{m+j-i}/s_{m-i}} = \prod_{0 \leq k \leq j-1} (R_{m+k} \cdots R_{m+k-(i-1)}) \qquad (m \geq i+1,\ i,\ j \geq 0). \tag{35}$$

Setting $j = i$ in (35) gives

$$\frac{s_{m-i}s_{m+i}}{s_m^2} = R_{m-(i-1)}R_{m-(i-2)}^2 \cdots R_m^i \cdots R_{m+(i-2)}^2 R_{m+(i-1)} \qquad (m \geq i+1,\ i \geq 0) \tag{36}$$

and setting $j = i+1$ gives

$$\frac{s_{m-i}s_{m+i+1}}{s_m s_{m+1}} = R_{m-(i-1)}R_{m-(i-2)}^2 \cdots R_m^i R_{m+1}^i \cdots R_{m+i-1}^2 R_{m+i} \qquad (m \geq i+1,\ i \geq 0). \tag{37}$$

Now, recalling that

$$s_n = \beta_n(s_1 s_{n-1} + s_2 s_{n-2} + \cdots + s_{n-1}s_1) \qquad (n \geq 2),$$

we get, for $m \geq 3$,

$$R_{2m} = \frac{\beta_{2m-1}\beta_{2m+1}}{\beta_{2m}^2} \cdot \frac{(2s_{m-1}s_m + 2s_{m-2}s_{m+1} + \sum_{2 \leq i \leq m-2} 2s_{m-i-1}s_{m+i})(2s_m s_{m+1} + 2s_{m-1}s_{m+2} + \sum_{2 \leq i \leq m-1} 2s_{m-i}s_{m+i+1})}{(s_m^2 + 2s_{m-1}s_{m+1} + \sum_{2 \leq i \leq m-1} 2s_{m-i}s_{m+i})^2},$$

so

$$R_{2m} = 4\frac{\beta_{2m-1}\beta_{2m+1}}{\beta_{2m}^2} R_m \cdot \frac{(1 + R_{m-1}R_m + \sum_{2 \leq i \leq m-2} R_{m-i} \cdots R_{m-1}^i R_m^i \cdots R_{m+(i-1)}) \cdot (1 + R_m R_{m+1} + \sum_{2 \leq i \leq m-1} R_{m-(i-1)} \cdots R_m^i R_{m+1}^i \cdots R_{m+i})}{(1 + 2R_m + \sum_{2 \leq i \leq m-1} 2R_{m-(i-1)} \cdots R_m^i \cdots R_{m+(i-1)})^2}. \tag{38}$$

Similarly, for $m \geq 2$,

$$R_{2m+1} = \frac{\beta_{2m}\beta_{2m+2}}{\beta_{2m+1}^2} \cdot \frac{(s_m^2 + 2s_{m-1}s_{m+1} + \sum_{2 \leq i \leq m-1} 2s_{m-i}s_{m+i})(s_{m+1}^2 + 2s_m s_{m+2} + \sum_{2 \leq i \leq m} 2s_{m+1-i}s_{m+1+i})}{(2s_m s_{m+1} + 2s_{m-1}s_{m+2} + \sum_{2 \leq i \leq m-1} 2s_{m-i}s_{m+i+1})^2},$$

so

$$R_{2m+1} = \frac{1}{4}\frac{\beta_{2m}\beta_{2m+2}}{\beta_{2m+1}^2}\cdot \tag{39}$$

$$\frac{\begin{array}{c}(1 + 2R_m + \sum\limits_{2\le i\le m-1} 2R_{m-(i-1)}\cdots R_m^i \cdots R_{m+(i-1)})\cdot \\ (1 + 2R_{m+1} + \sum\limits_{2\le i\le m} 2R_{m+1-(i-1)}\cdots R_{m+1}^i \cdots R_{m+1+(i-1)})\end{array}}{(1 + R_m R_{m+1} + \sum\limits_{2\le i\le m-1} R_{m-(i-1)}\cdots R_m^i R_{m+1}^i \cdots R_{m+i})^2}.$$

We now prove by induction that $R_n \le R_+ = \frac{1}{2}$ for all n. For $2 \le n \le 4$ this can be proven by direct computation. Otherwise, fix some $n \ge 5$, and assume that R_2, $\ldots$, $R_{n-1} \le \frac{1}{2}$. Set

$$\xi_n = \frac{\beta_{n-1}\beta_{n+1}}{\beta_n^2}.$$

It is easy to show that (since $n \ge 5$) $\xi_n \le \frac{10}{19}$. Now first suppose that n is even, so $n = 2m$, $m \ge 3$. Then from (38), we immediately have

$$\frac{R_{2m}}{4\xi_{2m}} \le R_m \frac{(1 + R_+ R_m + \sum_{i\ge 2} R_+^{i(i+1)})^2}{(1 + 2R_m)^2} = A, \qquad \text{say.}$$

We will prove that $A \le \frac{19}{80}$; if this is so, then

$$R_{2m} \le 4\xi_{2m}\frac{19}{80} \le 4\cdot\frac{10}{19}\cdot\frac{19}{80} = \frac{1}{2},$$

as desired. However, $\sum_{i\ge 2} R_+^{i(i+1)} \le 0.1$, so we will have $A \le \frac{19}{80}$ provided that

$$(\frac{1}{2}R_m + 1.1)^2 R_m \le \frac{19}{80}(1 + 2R_m)^2,$$

i.e., if

$$\frac{1}{4}R_m^3 + 1.1R_m^2 + 1.21R_m \le \frac{19}{80} + \frac{19}{20}R_m + \frac{19}{20}R_m^2.$$

Since $R_m \le R_+ = \frac{1}{2}$, it will do to have

$$(\frac{19}{80} - \frac{1}{4}\cdot(\frac{1}{2})^3) + (\frac{19}{20} - 1.21)R_m + (\frac{19}{20} - 1.1)R_m^2 \ge 0,$$

which is true as the expression is positive for $R_m = \frac{1}{2}$ and nonincreasing for $R_m \in [0, \frac{1}{2}]$.

The other possibility is that $n = 2m+1$ is odd and $m \ge 2$. Then from (39),

$$\frac{R_{2m+1}}{\frac{1}{4}\xi_{2m+1}} \le \frac{(1 + 2R_m + \sum_{i\ge 2} 2R_+^{i^2})(1 + 2R_{m+1} + \sum_{i\ge 2} 2R_+^{i^2})}{(1 + R_m R_{m+1})^2} = B, \qquad \text{say.}$$

Since $\sum_{i\geq 2} 2R_+^{i^2} \leq 0.13$, we have

$$\begin{aligned} B &\leq \frac{(1.13+2R_m)(1.13+2R_{m+1})}{1+2R_mR_{m+1}} \\ &= \frac{1.13^2+2.26(R_m+R_{m+1})+4R_mR_{m+1}}{1+2R_mR_{m+1}} \\ &\leq 1.13^2+2.26+(4-2\cdot 1.13^2)\frac{R_mR_{m+1}}{1+2R_mR_{m+1}} \\ &\leq 1.13^2+2.26+(4-2\cdot 1.13^2)\frac{1/4}{1+2/4} \\ &= 1.13^2+2.26+(4-2\cdot 1.13^2)\frac{1}{6} \leq 3.8, \end{aligned}$$

so

$$R_{2m+1} \leq 3.8\frac{1}{4}\xi_{2m+1} \leq 3.8\cdot\frac{1}{4}\cdot\frac{10}{19} = \frac{1}{2},$$

as desired. □

Theorem 9.2. *Suppose that $q = 2$. Then if $n \geq 2$, $n = qx + y$, and $0 \leq y \leq q$, we have*

$$\sum_{|b|=n}{}' \prod_{i=0}^{q-1} s_{b_i} \bigg/ s_x^{q-y}s_{x+1}^y \leq 3. \tag{40}$$

Proof. Using (34), (36), (37) and Lemma 9.1, we find that

$$s_{x-i}s_{x+i} \leq 2^{-i^2}s_x^2 \qquad (x \geq i+1,\ i \geq 0)$$

and

$$s_{x-i}s_{x+i+1} \leq 2^{-i(i+1)}s_xs_{x+1} \qquad (x \geq i+1,\ i \geq 0).$$

It follows that, if $n = 2x$ is even,

$$\sum_{j+k=n,\ j,k>0} s_js_k \leq s_x^2(1+2\sum_{i>0}2^{-i^2}) \leq 3s_x^2,$$

and if $n = 2x+1$ is odd,

$$\sum_{j+k=n,\ j,k>0} s_js_k \leq s_xs_{x+1}\cdot 2\sum_{i\geq 0}2^{-i(i+1)} \leq 3s_xs_{x+1}.$$

This completes the proof. □

If $q = 2$, we therefore have

$$\begin{aligned} \log_q s_n &= \log_q \beta_n + (q-y)\log_q s_x + y\log_q s_{x+1} + O(1), \\ &\qquad \text{where } n \geq 2,\ n = qx+y,\ 0 \leq y \leq q. \end{aligned} \tag{41}$$

We conjecture that such a recursion also holds when $q > 2$ is integral. We now show how to reduce recursions like those above to the simpler recursion of Section 8.

Lemma 9.3. *Fix q and some constant $\bar{C}$, and for nonnegative integers n, set*

$$\begin{aligned} w_n &= -\tfrac{n^2}{2(1-q^{-1})} - \tfrac{1}{2}n\log_q n, && \textit{if } n > 0, \\ w_n &= 0, && \textit{if } n = 0. \end{aligned}$$

Then there is some constant C' such that if $a_1, \ldots, a_q$ satisfy $a_1+\cdots+a_q = n \geq 1$, and for $i = 1, \ldots, q$, we have $|a_i - n/q| < \bar{C}$, then

$$|w_n - \log_q \beta_n - (w_{a_1} + \cdots + w_{a_q})| < C'. \tag{42}$$

Proof. Write $a_i = n/q + \epsilon_i$. It will suffice to prove (42) for large n. Take n large enough so that $|\epsilon_i| < \bar{C} < n/2q$. Now

$$\begin{aligned} w_{a_i} &= -\frac{(n/q+\epsilon_i)^2}{2(1-q^{-1})} - \frac{1}{2}(\frac{n}{q}+\epsilon_i)\log_q(\frac{n}{q}+\epsilon_i) \\ &= -\frac{n^2}{2(q^2-q)} - \epsilon_i\frac{n/q}{1-q^{-1}} - \frac{\epsilon_i^2}{2(1-q^{-1})} \\ &\qquad -\frac{1}{2}(\frac{n}{q}+\epsilon_i)\log_q\frac{n}{q} - \frac{1}{2}(\frac{n}{q}+\epsilon_i)\log_q(1+\frac{\epsilon_i}{n/q}). \end{aligned}$$

Summing over i, we get

$$\begin{aligned} \sum_{1\leq i\leq q} w_{a_i} &= -\frac{n^2}{2(q-1)} - \frac{1}{2}n(-1+\log_q n) \\ &\qquad - \sum_{1\leq i\leq q}\frac{\epsilon_i^2}{2(1-q^{-1})} + \frac{1}{2}(\frac{n}{q}+\epsilon_i)\log_q(1+\frac{\epsilon_i}{n/q}). \end{aligned} \tag{43}$$

By looking at the power series for $\log(1+x)$, we find that $|\log_q(1+\chi)| \leq 4|\chi|$, if $|\chi| < \frac{1}{2}$. By our assumption on n, $|\epsilon_i|/(n/q) < \frac{1}{2}$, so $|\log_q(1+\epsilon_i/(n/q))| \leq 4|\epsilon_i|/(n/q)$. It follows that the absolute value of the sum on the right-hand side of (43) is bounded, say by C'', so

$$\left|-\frac{n^2}{2(q-1)} + \frac{1}{2}n - \frac{1}{2}n\log_q n - (w_{a_1}+\cdots+w_{a_q})\right| \leq C''.$$

The result now follows from the definition of w_n and part (i) of Lemma 7.8. □

The significance of Lemma 9.3 is that we can, given a recursion like (33) or (41), subtract off w_n from the unknown to obtain a recursion of the form considered in Section 8, and then apply Theorem 8.1. Applying this to (41) in the $q = 2$ case gives us (19), and concludes the proof of Theorem 7.1.

If (41) were to hold when $q > 2$, applying Lemma 9.3 to (41) and then applying Theorem 8.1 would give the

Conjecture 9.4. *If $q > 2$ is integral, we have*

$$\log_q r_n = -\frac{n^2}{2(q-1)} - \frac{1}{2}n\log_q n + (\frac{1}{2} + \bar{W}_q(\log_q n))n + O(1), \tag{44}$$

for some continuous $\bar{W}_q$ of period 1.

Finally, we cannot resist remarking that applying Lemma 9.3 to (33) and then applying Theorem 8.1 to solve the resulting recursion of form (28) shows that

$$\log_q \nu_n = -\frac{n^2}{2(1-q^{-1})} - \frac{1}{2}n\log_q n + W_q(\log_q n)n + O(1),$$

for some continuous W_q of period 1.

This could be used to give an alternate proof of Theorem 7.1. Since ν_n is the greatest tree contribution to r_n, and since the $\log_q$ of the number of labelled q-trees with n leaves is $O(n)$ (Lemma 7.3), it follows that

$$\log_q r_n = -\frac{n^2}{2(q-1)} - \frac{1}{2} n \log_q n + O(n).$$

We show graphs of W_2 and $\bar{W}_2$ below.

References

[1] Auel, Asher N. Volumes of Integer Polynomials over Local Fields. B. A. thesis, Reed College, May 2003.

[2] Euler, L. Introductio in Analysin Infinitorum **1** (1748), §76.

Center for Communications Research, 4320 Westerra Court, San Diego, CA 92121

E-mail address: {buhler,dgoldste,dmoews,joelr}@ccrwest.org

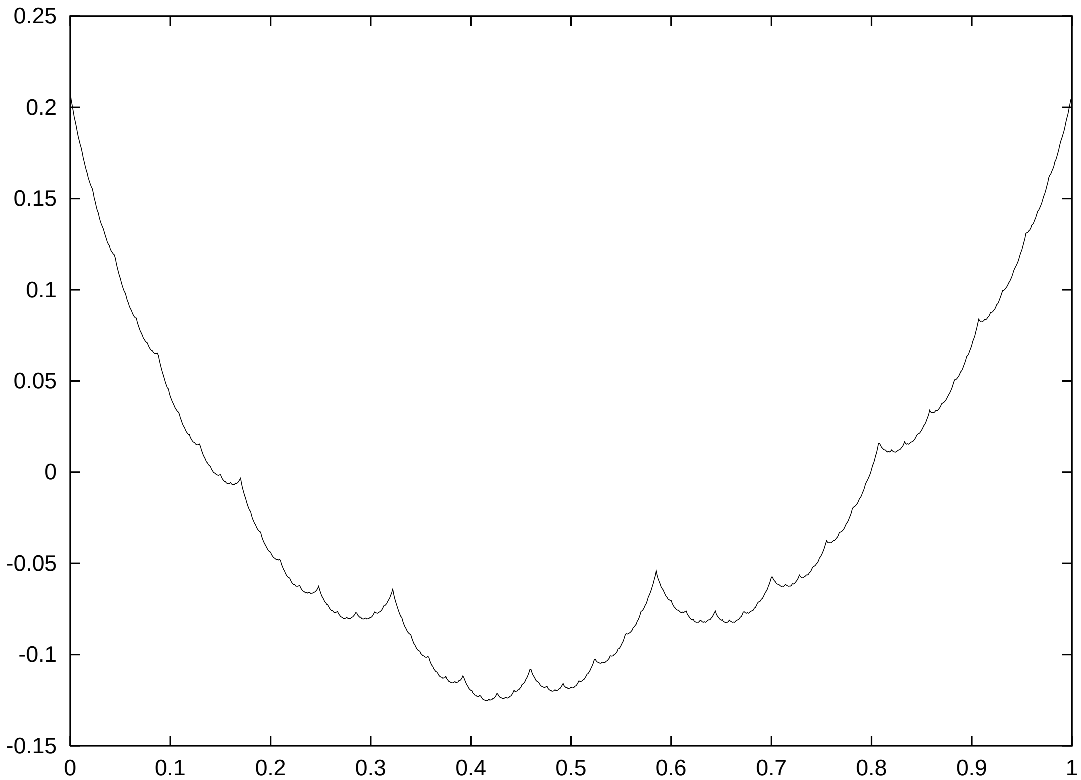

FIGURE 1. Graph of W_2 over its period.

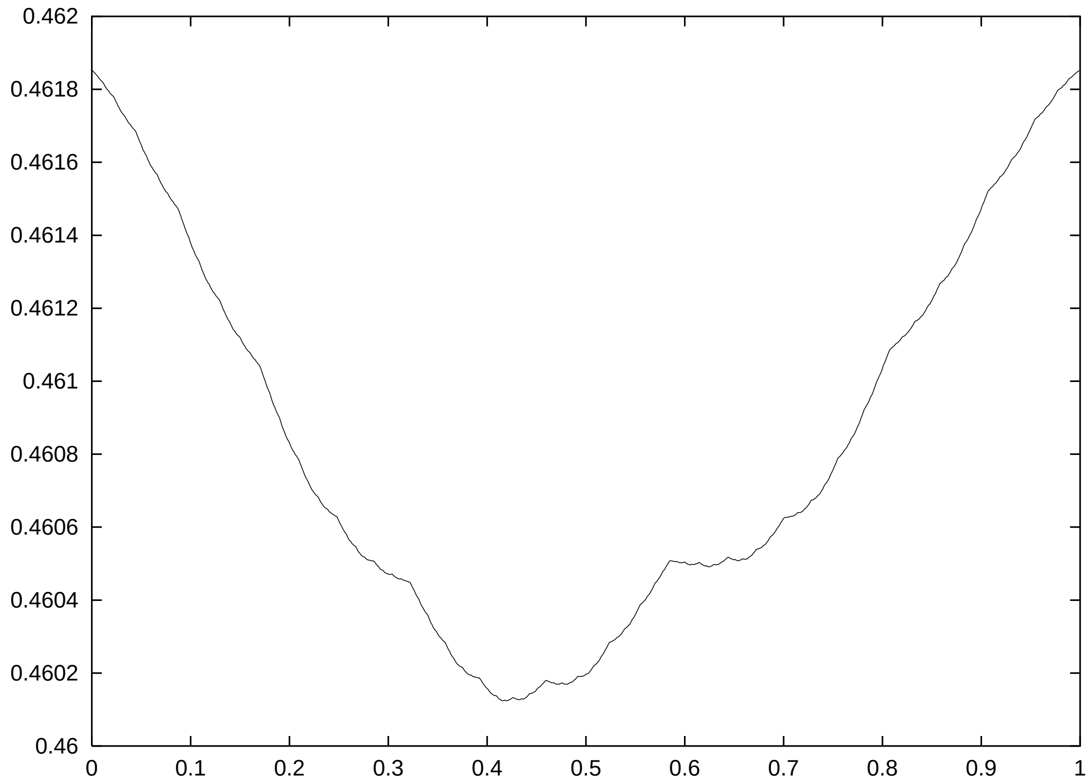

FIGURE 2. Graph of $\bar{W}_2$ over its period.